

课程名称：法律服务与司法管理技术

课 程 号：10001254

任课教师：周蔚

姓　　名：刘龙

学　　号：2201010146

评分区域（由阅卷老师红笔填写）：

总　分：________________

评分人：________________

大数据法律监督之研究

【摘要】

大数据法律监督是当下的热门话题，也是新一轮检察体制改革的关键所在。大数据法律监督对检察机关提出了新要求，需要对其基本理论有所把握。作为新生的事物，大数据法律监督仍然在蓬勃发展之中，难免会遇到各个方面的问题，其中最突出的、争议也是最多是数据、算法和制度三个方面。若要大数据法律监督发挥其预期作用，则需要从多个角度共同改进。本文针对目前存在的问题，根据我国的实际情况，提出了一些完善措施，为相关规定的出台提供些许参考。

【关键词】大数据 法律监督 算法 制度

一、 引言

习近平总书记指出“实践反复告诉我们，关键核心技术是要不来、买不来、讨不来的”，“中国要强盛、要复兴，就一定要大力发展科学技术”，“要把握数字化、网络化、智能化融合发展的契机，以信息化、智能化为杠杆培育新动能”，“以‘鼎新’带动‘革故’，以增量带动存量”。[①]检察机关作为我国国家机构的重要组成部分，承担着法律监督、惩治犯罪的重要职能，一方面，应当健全保护创新的法治环境，为科技创新提供更加完备的司法保障；另一方面也需要深化智慧检务建设，与时俱进，抢占先机，直面问题，迎难而上，肩负起历史赋予的重任。

2021 年 6 月 15 日，《中共中央关于加强新时代检察机关法律监督工作的意见》制定出台，明确提出了检察机关充分发挥法律监督职能加强智能化建设的要求。2022 年 1 月，最高人民检察院检察长张军在全国检察长（扩大）会议上强调，要增强大数据战略思维，运用大数据助力法律监督“本”的提升和“质”的嬗变。检察大数据战略的提出，是检察机关深入学习贯彻习近平法治思想和习近平总书记关

[①] 习近平在中国科学院第十九次院士大会、中国工程院第十四次院士大会上的讲话。

于网络强国、数字中国的重要论述，贯彻新发展理念、适应信息化时代新趋势的重要举措，充分彰显了检察机关满足现代社会治理需求的责任感和使命感，有利于促进强化法律监督手段，提升法律监督实效，加快推进检察工作质量变革、效率变革以及动力变革，助推国家治理体系和治理能力现代化。

贯彻落实好检察大数据战略，要在强化思想认识、拓展信息来源、强化业务培训和综合应用上下真功夫、深功夫，为促进检察工作高质量发展，推进国家治理体系和治理能力现代化贡献更大检察力量。随着信息技术飞速发展和国家大数据战略深入实施，大数据在经济社会高质量发展中的作用愈加重要。落实全面依法治国，推进国家治理体系和治理能力现代化，离不开大数据。习近平总书记多次强调，要运用大数据提升国家治理现代化水平。关于大数据赋能法律监督，张军检察长曾作过一个生动的比喻:“监督办案就好像农耕,不掌握春种夏锄秋收传统知识、基本农技不行,但信息化时代,仅有这些传统的、基础的知识、技术已远远不够， 必须用科技、大数据手段提升质效才可能提高‘产能’！”如何提升检察人员深度应用大数据的能力，实现精准监督？如何运用大数据推动法律监督在抓好“本”的提升的基础上实现“质”的嬗变？如何促进解决社会治理深层次的问题？这些成为法律人面前不可回避的问题。

二、 大数据法律监督的基本理论

（一） 概念辨析

理论生成的一项基本任务是对特定事物概念进行合理界定。作为一个新生事物，大数据法律监督的内涵和外延其实在理论上并不是那么明确，故而，对其进行概念辨析不可忽视。但是实践中仍然有许多检察机关存在含混：或者将其与辅助办案混为一谈，如投诉举报查询、资金流向分析、电子证据勘查等辅助办案工具，主要用于搜集线索，与大数据法律监督并不完全一致；或者将其与事后的数据统计混为一谈，如单纯统计数据用于专题调研或趋势预测，只是单纯的记录数据，并不能从数据中发现未知线索；或者将其与数字案管混为一谈，如通过电子卷宗共享、分析办案数据，发现案件遗漏、错误，却也只是注重看到了技术层面的便捷高效，并不是需要重点进行改革的大数据法

律监督。[2]

简单来说，大数据法律监督应当是大数据驱动下的法律监督，大数据为法律监督注入新的动力。而大数据的使用其实在建国之初便有所体现，在选定国庆节的具体日期时，工作人员评估了全年的天气数据，发现在往年的大数据中，10 月 1 日的晴天数据较多，因此选定 10 月 1 日作为国庆节。而将大数据运用到法律监督之上时，用简单的公式表达就是“法律监督 + 大数据”。其中，法律监督指的是“专门的国家机关根据法律的授权，运用法律规定的手段对法律实施情况进行监察、督促并能产生法定效力的专门工作”，[3]是我国检察机关依法履行宪法法律赋予职责的特殊监督；大数据可以被简单理解为“为决策问题提供服务的大数据集、大数据技术和大数据应用的总称”。[4]此外，仍需注意的是，与国外不同，大数据法律监督的相关探索在我国呈现为“由国家力量主导”的一种特殊样态。[5]

（二） 基本特点

大数据法律监督是从数据或个案中发现规律与问题，通过总结、归纳并开发监督模型，从海量数据中挖掘隐含在其中的、仅依据个案难以知晓的监督信息和类案线索，继而从类案问题中分析、发现管理衔接、制度机制、立法执法司法等方面存在的系统性漏洞，以引导侦查、提出检察建议等方式，促进社会治理。具体来说，大数据法律监督有以下四个特点：[6]一是从个案到类案再到社会治理。大数据法律监督的路径是从个案中总结规律、特征，后根据规律、特征在海量数据中筛查出类案，在批量类案中发现立法、执法、司法、机制等方面存在的漏洞，最终落脚点是解决社会治理问题。二是精准发现类案线索。大数据法律监督需要从具体的监督场景或者小切口着手，提炼出有效的类案特征和正确的逻辑规律，最终从海量数据中精准识别类案监督线索。三是监督目标具有普遍性。大数据法律监督针对的监督对象或漏洞问题具有广泛性、普遍性，在现实中存在批量类案，且在其他地方或领域也同样存在，具有社会治理意义。四是技术上具有可行性。大数据法律监督在技术上要现实、可行，在实践中才能被应用、

---

[2] 翁跃强,文晓晴,姜琪等.大数据赋能法律监督的价值与应用[J].人民检察,2022(11):27-30.
[3] 张智辉.法律监督三辨析[J].中国法学,2003(05):14-22.DOI:10.14111/j.cnki.zgfx.2003.05.002.
[4] 朱扬勇,熊赟.大数据是数据、技术,还是应用[J].大数据,2015,1(01):78-88.
[5] 刘品新.智慧司法的中国创新[J].国家检察官学院学报,2021,29(03):81-101.
[6] 翁跃强,文晓晴,姜琪等.大数据赋能法律监督的价值与应用[J].人民检察,2022(11):27-30.

复制、推广。数据要可采集、可转化、可清洗，能通过筛查、碰撞、聚类、统计等具体方法挖掘出数据背后的隐含信息。

（三） 重要意义

作为法律监督的一轮重塑变革，大数据法律监督巧妙构造了一种阿基米德杠杆模型。阿基米德发现，借助杠杆可以达到无限省力的效果。展开来说，作用在杠杆两端的力与力臂的乘积相等，杠杆可平衡。杠杆原理的公式化表达是，动力×动力臂=阻力×阻力臂。该原理表明，法治建设中选择的支点，越扩展动力及动力臂一端，就越能达到目标；越压缩阻力及阻力臂一端，效果亦然。大数据法律监督助推法治建设的奥秘即在于此。

其一，大数据法律监督扩增了检察官甄别违法案件的“动能”。传统上，法律监督的启动源于当事人举报等途径的个案线索，检察官启动调查核实，线索单一、渠道受限。大数据法律监督为检察官插上了大数据分析与数据模型、平台的翅膀，开辟了类案监督、智能监管的新天空。如湖北省检察院政务数据云监督平台能够分析比对全省百万级、千万级的行政执法案件数据，浙江省绍兴市检察院虚假诉讼类监督模型能够探测北京等外地的虚假诉讼案件线索，均是难以想象的飞跃。同时，大数据法律监督可以对异常案件“画像”且转化为智能化预测预警条件，达到“上医治未病”的效果。

其二，大数据法律监督延长了检察官治理违法行为的“动力臂”。我国检察系统长期存在监督与办案“两张皮”现象，大数据法律监督将法律监督紧密融入“四大检察”“十大业务”，实现业务部门同案管部门、技术部门等良性联动，催生出新型融合式监督。更重要的是，大数据法律监督实行“一轴多元”共治架构，由检察机关主导，会同公安机关、审判机关、司法行政机关、行政机关以及社会组织、行业协会等主体，通过促成多方合力，实现真正“双赢多赢共赢”。

其三，大数据法律监督压缩了违法案件的存在空间和破坏力。这主要表现为其分解了违法对象的辐射面、“朋友圈”，遏制了违法行为延续的不良影响。所体现出的治理优势，一是智慧锁定各种批量异常案件进行多案纠错，二是解析社会治理惯常症结进行类案防错。如江苏省苏州市检察院“破产领域虚假劳资债权监督模型”，已经筛选出风险企业涉劳动纠纷成百上千条线索，对发现涉劳动报酬纠纷的批量违法案件均全面纠正。之后，该院深化信息共享等机制，构建了长效

治理体系。这就产生了值得赞许的治理效能。

三、 大数据法律监督的隐忧

有学者指出，电子政务项目风险涵盖电子政务项目的各要素，如项目范围风险、项目成本风险、项目进度风险、项目质量风险、项目团队风险等，并且各类项目风险相互关联、相互叠加。[⑦]大数据法律监督作为电子政务的重要组成部分，也不能免俗。在看到大数据法律监督的重要意义的同时，也应当保持清醒的头脑，对可能存在的风险作出及时的防范，下文将简要分析大数据法律监督的隐忧与风险。

“统治社会对应的是简单或力学的机器，规训社会对应的是高能的机器；控制社会对应的是控制学或电脑……面对出现在开放环境中的那些不间断的控制形式，可能最严酷的禁锢对我们来说仿佛是美妙的回忆”。[⑧]德勒兹的担忧不无道理，尤其是当大数据智能的发展已经超出人们想象的时候，人们可能感到更多的是恐惧而不是欣喜。从数据、算法和制度三个方面进行简单梳理，大数据法律监督存在以下几方面的隐忧。

（一） 数据方面

数据可以说是大数据法律监督的基础，是万里长城中的一块砖。但是千里之堤往往溃于蚁穴，在数据方面，可能会存在以下几种风险，严重的时候可能影响检察权的公正行使：第一，数据可能不全面。全面是大数据法律监督的基本要求之一，数据量的全面与否，直接决定分析结果的准确性。大数据司法监督中的证据审查、量刑偏离度分析等机制过度依赖“中国裁判文书网”的数据。然而，网上公开的3000多万份裁判文书实际只占总量的50%左右，还有约一半的裁判文书并未在网上公布。[⑨]除终局性裁判文书外，而一般如侦查终结报告、起诉意见书等过程性文书的数据化程序也不够。第二，数据可能不真实。数据本身可能会存在错误，而依据错误的数据得出错误的结论是不可

---

⑦ 赵志刚，金鸿浩著：《智慧检务概论：检察机关法律监督的科技智慧》，中国检察出版社2018年版，第363页。

⑧ [法] 吉尔·德勒兹：《哲学与权力的谈判：德勒兹访谈录》，刘汉全译，商务印书馆 2000 年版，第 200 页。

⑨ 左卫民.关于法律人工智能在中国运用前景的若干思考[J].清华法学,2018,12(02):108-124.

避免的。具体到大数据法律监督而言，如何保证掺杂了许多主观因素的判断下，许多数据被正确地识别，成为了重中之重。比如不同判决对于类似行为作出的不同认定，这些偏差、错误的裁判数据带来了“数据噪音”，相当于以错误的司法经验、裁判规则去指导个案，在此基础上所产生的算法模型必然会带有一定的偏差。第三，数据可能不可靠。这一点其实是前面两点的叠加态，更进一步加剧了数据对于大数据法律监督的影响。不难想象，在这个数据爆炸的时代，每天都会有新的数据产生，数据的更迭令人应接不暇。对于大数据法律监督而言，司法领域每天都产生海量新的裁判文书，新数据的不断涌入意味着其中可能蕴含着新的裁判规则和监督规则。新的司法数据带来了新的裁判经验和监督规则，从而导致既有的司法监督模型不够全面、不够精细。并且，随着大数据技术发展的迅速，会造成既有司法监督模型愈发滞后于新型司法权运行实务。

（二） 算法方面

如果说数据是大数据法律监督的砖，那么算法就是砌成万里长城的方法，光有了原材料，没有了制造流程也是行不通的。在算法方面，可能会存在以下几种风险，也是学界一直在讨论的：第一，算法可能不智能。大数据司法监督的算法产生主要有“自底向上”和“自顶向下”两种方式。前者是指借助一定的技术手段，从海量数据中提取出资源模式，由机器对海量数据学习后自动生成算法模型；后者是指借助既定的司法规则、司法经验，人为构建算法规则。[⑩]目前，鉴于人工智能发展水平的限度以及司法专业本身的难度，大部分算法生成仍采取“自顶向下”的方法，将主观司法经验进行算法规则翻译，所谓“有多少人工就有多少智能”，大数据司法监督的算法产生过程仍然不够自动化、智能化。现实法律世界存在着大量未经经验浸润的法律逻辑，也存在着未经逻辑规整的经验。[⑪]这就有可能导致罪名覆盖不完整，对应不精确等问题。第二，算法可能不公正。不管是对于监督者还是对于被监督者来说，算法其实都是不透明的，往往会形成所谓的“黑箱效应”。在给检察官带来便利，省去了学习成本的同时，应当清醒地认识到，算法的不公开极有可能为大数据法律监督蒙上一层不公正的面纱，降低群众的信赖程度，甚至什么数据会被采集、参数

[⑩] 刘峤,李杨,段宏等.知识图谱构建技术综述[J].计算机研究与发展,2016,53(03):582-600.
[⑪] 郑戈.算法的法律与法律的算法[J].中国法律评论,2018(02):66-85.

是否正确这类技术性的问题恐怕也难以由检察官来进行核验。这会导致司法监督过程的不透明，进而带来司法监督程序的不公正。第三，算法可能存在差异化。这似乎是业界一直以来担忧的问题之一，315晚会上披露的“大数据杀熟”历历在目。我国幅员辽阔，不同地区固然存在差异，那么标准如何制定，差异化的算法是否会进一步导致案件遭受到不平等的对待，值得进行进一步的研究。

（三） 制度方面

相对而言，制度方面的风险较为宏观，关涉到观念层面的问题。第一，任何权力都存在滥用的风险，大数据法律监督应当有所限度，滥用可能会导致检察权的无序扩张，也会影响被监督者的积极性，产生恐惧和抵触心理，甚至采取反制手段，规避大数据法律监督。长此以往，将压缩司法权的运行范围，僵化司法权的运行。 第二，机械式的裁判导致思维固化，各地法院愈发成为“判决工厂”，产生算法支配个案司法的情景面对法律争议，司法人员不再诉诸严格的对抗式审判、法律推理等程序性要求，而是诉诸机器判断结果。[12]司法裁判的本质是价值判断，而算法本质是数据运算，其无法像司法官一样进行带有感情色彩的价值判断。[13]

四、 相关隐忧的应对措施

通过上文的分析，不难发现，大数据法律监督的隐忧存在四个特点：一是承受程度低，一旦发生重大问题，其后果难以估量；二是突发几率高，涉及到全国上下四级三千余检察机关和数十万工作人员；三是控制过程较长，需要一支稳定、高素质的工作团队；四是要素类别多样，必须贯穿于各阶段。

其一，加强党对大数据法律监督的领导。正是基于我国政治体制的这一特点， 中央党政部门发布统一的标准和规定,自上而下地推行检察监督智能化及其具体要求,这将使大数据法律监督的改革的落实

---

[12] 季卫东.人工智能时代的司法权之变[J].东方法学,2018(01):125-133.DOI:10.19404/j.cnki.dffx.2018.01.013.

[13] 王燃.大数据司法监督机制研究[J].湖南科技大学学报(社会科学版),2021,24(03):132-141.DOI:10.13582/j.cnki.1672-7835.2021.03.017.

具有可行性。[14]

其二，加强对大数据法律监督的数据使用。加强对数据的收集、分析和研判，实现互联互通、信息共享共用，依法有力督促、引导司法机关全面调取电子数据等客观证据，形成完整的证据链条，为顺利办案奠定坚实的基础，提高精准办案质量。[15]

其三，加强对大数据法律监督的算法控制。大数据法律监督机制应具有可验证性与透明性，监督的数据来源、算法原理都应当向被监督对象公开，从而保障司法监督的程序正当。实际上，算法可解释性乃至可视化是可以用技术解决的问题，一旦法律提出了相关要求，技术界便会想方设法使算法成为可解释。[16]

其四，加强对大数据法律监督的制度设计。明确大数据法律监督的责任链条，明确区分不同主体之间的责任，例如机器与技术人员的责任、系统部署者的责任、司法官的责任等等，通过明确的制度将一系列要求落到实处。充分、合理地运用人工智能、大数据等数字技术，顺应和借助司法数字化的趋势，推进我国司法责任体系的建设和完善。[17]

其五，建立完善对大数据法律监督的评估机制。每年由各级政法委会同司法机关，对本级大数据法律监督进行全面客观的评价，并公布评价报告，监督整改，畅通反馈途径，积极倾听群众声音。

五、 结语

英国学者理查德 · 萨斯坎德在《法律人的明天会怎样？》一书中，告诫法律人要睿智地思考未来。[18]未来已来，将至已至，既见未来，为何不备？相较于认识停留于过去与现实，预测人类世界的走向，是更广泛而有意义的挑战。于大数据法律监督的探索而言，这就要求检察人群体要勇做行动派，实现数字检察转轨，依托大数据技术铺开多案监督。于国家、社会治理现代化的推进而言，检察机关同有关机关要拓展“同心圆”，以法律监督之多赢双赢共赢方式，协同建设“中

[14] 赵毅宇.检察监督智能化的发展隐忧及应对逻辑[J].法制与社会发展,2023,29(02):178-194.
[15] 张俊杰,马晓怡,刘德印.大数据+检察监督管理运行创新机制探讨[J].中国检察官,2018(19):25-28.
[16] 郑戈.算法的法律与法律的算法[J].中国法律评论,2018(02):66-85.
[17] 高童非.数字时代司法责任伦理之守正[J].法制与社会发展,2022,28(01):151-172.
[18] [英] 理查德 · 萨斯坎德:《法律人的明天会怎样？——法律职业的未来》，何广越译，北京大学出版社 2015 年版，第 7-8 页。

国之治”。[19]制度的一系列过程都逐渐积累于经过的岁月长河中，沉淀产生在人民群众的历史推动中。[20]大数据法律监督仍在蓬勃发展之中，需要理性辩证地去看待新的制度形成和演变，也需要留待实践和时间进行检验。

参考文献

著作类：

[1]赵志刚，金鸿浩著：《智慧检务概论：检察机关法律监督的科技智慧》，中国检察出版社 2018 年版。

[2][法] 吉尔·德勒兹: 《哲学与权力的谈判: 德勒兹访谈录》, 刘汉全译, 商务印书馆 2000 年版。

[3]［英］理查德 · 萨斯坎德:《法律人的明天会怎样？——法律职业的未来》，何广越译，北京大学出版社 2015 年版。

论文类：

[1]翁跃强,文晓晴,姜琪等.大数据赋能法律监督的价值与应用[J].人民检察,2022(11):27-30.

[2]张智辉.法律监督三辨析[J].中国法学,2003(05):14-22.DOI:10.14111/j.cnki.zgfx.2003.05.002.

[3]朱扬勇,熊赟.大数据是数据、技术,还是应用[J].大数据,2015,1(01):78-88.

[4]刘品新.智慧司法的中国创新[J].国家检察官学院学报,2021,29(03):81-101.

[5]翁跃强,文晓晴,姜琪等.大数据赋能法律监督的价值与应用[J].人民检察,2022(11):27-30.

[6]左卫民.关于法律人工智能在中国运用前景的若干思考[J].清华法学,2018,12(02):108-124.

[7]刘峤,李杨,段宏等.知识图谱构建技术综述[J].计算机研究与发展,2016,53(03):582-600.

[8]郑戈.算法的法律与法律的算法[J].中国法律评论,2018(02):66-85.

[9]季卫东.人工智能时代的司法权之变[J].东方法学,2018(01):125-133.DOI:10.19404/j.cnki.dffx.2018.01.013.

[10]王燃.大数据司法监督机制研究[J].湖南科技大学学报(社会科学版),2021,24(03):132-141.DOI:10.13582/j.cnki.1672-7835.2021.03.017.

[19] 刘品新.论大数据法律监督[J].国家检察官学院学报,2023,31(01):76-92.
[20] 朱苏力.制度是如何形成的?——关于马歇尔诉麦迪逊案的故事[J].比较法研究,1998(01):68-74.

[11]赵毅宇.检察监督智能化的发展隐忧及应对逻辑[J].法制与社会发展,2023,29(02):178-194.
[12]张俊杰,马晓怡,刘德印.大数据+检察监督管理运行创新机制探讨[J].中国检察官,2018(19):25-28.
[13]郑戈.算法的法律与法律的算法[J].中国法律评论,2018(02):66-85.
[14]高童非.数字时代司法责任伦理之守正[J].法制与社会发展,2022,28(01):151-172.
[15]刘品新.论大数据法律监督[J].国家检察官学院学报,2023,31(01):76-92.
[16]朱苏力.制度是如何形成的?——关于马歇尔诉麦迪逊案的故事[J].比较法研究,1998(01):68-74.